Journal of Life Sciences & Biomedicine, vol. 3(76), No 1, p. 60-66 (2021)

In silico analysis of Dreb transcription factor genes in bread wheat

S.M. Rustamova

Institute of Molecular Biology & Biotechnologies, Azerbaijan National Academy of Sciences, 11 Izzat Nabiyev Str., Baku AZ 1073, Azerbaijan
***For correspondence**: s.rustamova@imbb.science.az*

Received: March 09, 2021; Received in revised form: April 09, 2021; Accepted: May 09, 2021

The dehydration responsive element binding (Dreb) genes which are representatives of the AP2 / ERF transcription factor family and playing a key role in drought-induced transcriptome in wheat were studied using *in silico* methods. For this purpose, information on relevant genes (Accession Nr. AF303376.1, AB193608.1, KM520370.1, DQ195068.1) was obtained from the NCBI. FASTA data of proteins corresponding to each gene were analyzed comparatively by the MAFFT CLUSTAL format alignment software, and the main conservative areas were identified. Two conservative functional amino acids specific for AP2 domain - valine and glutamine - were identified in positions 14 and 19 in all studied genes. Specific amino acid substitutions have been identified in the protein (DQ195068.1) that binds to the dehydration element specific to the D genome in the areas involved in the formation of the nuclear localization signal (NLS) and the α-helix structure. The results obtained could be a scientific basis for future laboratory studies of Dreb genes in wheat.

Keywords: *Dreb, AP2 domen, nuclear localization signal (NLS), α-helix, in silico analysis*

INTRODUCTION

Although the advances in genomics contributed to the improvement of some agriculturally important crops, similar efforts in wheat (Triticum spp.) were more challenging. This is attributed to the size and complexity of the wheat genome, and the lack of genome-assembly data for multiple wheat lines (Walkowiak et al., 2020, Alotaibi et al., 2021). The current knowledge of wheat biology and the molecular basis of central agronomic traits are not sufficient for wheat breeding. There is an urgent need for wheat research and breeding to accelerate genetic gain as well as to increase and protect wheat yield and quality traits for meeting the demands of human population growth (Zhu et al., 2021). Clarification of gene functions and availability of wheat genome sequence information as well as genome editing methods will open up new opportunities for improving crops under stress conditions (Rathan, 2021).

Some candidate genes involved in the adaptive responses to abiotic stress have been determined in cereals. Transcription regulators have been found to play an important role in the adaptation of plants against changing environmental conditions. According to large-scale transcriptome analyses, protective proteins and regulatory proteins are the main types of molecular stress responses (Shahzad et al., 2021). Proteins such as chaperones, osmotin, antifreeze proteins, mRNA-binding proteins protect cells against stress conditions. Regulatory proteins such as transcription factors, including myeloblastosis oncogene (MYB), basic leucine zipper (bZIP), NAM, ATAF, CUC (NAC), and dehydration responsive element binding (DREB) proteins rearrange the gene expressions to protect the plant from stresses. The interaction between transcription factors and cis-elements of target gene promoters is very important for the regulation of stress-related gene expression.

Dehydration responsive element-binding proteins are essential transcription factors that stimulate stress-related genes (Niu et al., 2020). Two types of DREB transcription factors were

https://doi.org/10.29228/jlsb.8
Available online 30 June 2021

observed: DREB1 and DREB2. They are contained in different signal transduction pathways under low temperature and dehydration, respectively. The C-repeat (CRT) and low-temperature-responsive element (LTRE) known as cis-acting elements both contain an A/GCCGAC motif similar to the core of the DRE sequence regulating cold-inducible gene expression. DREB1/DREB2 homologous genes were identified in several grasses, including wheat, rice, barley, sorghum, maize, oat, rye, and perennial ryegrass. Dreb1/CBF genes are stimulated by cold, whereas the Dreb2 genes are generally stimulated by dehydration, high salinity, and heat (Hassan et al., 2021). DREB family proteins belonging to the AP2/ERF transcription factor family contain one AP2/ERF DNA-binding domain.

We aimed at the detailed analyses of Dreb genes in the bread wheat genome. For this purpose, we performed *in silico* comprehensive analysis of four Dreb genes using the available nucleotide and protein sequences from the current databases.

MATERIALS AND METHODS

Sequence Sources

The complete cDNA and corresponding protein sequences of DREB in wheat (*Triticum aestivum* L.) were retrieved from the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/).

Multiple Sequence Alignment

Multiple Sequence Alignment is generally the alignment of three or more biological sequences (protein or nucleic acid) of similar length. Based on the results, homology can be assessed and the evolutionary relationships between the sequences studied. MAFFT v7.427 (https://mafft.cbrc.jp/alignment/server/) CLUSTAL format alignment program was used for sequencing. MAFFT (Multiple Alignment using Fast Fourier Transform) is a high-speed multiple sequence alignment program.

Annotation of Functional Motifs

For searching functional motives in the selected Dreb genes, the relevant functions of Softberry, Inc. software (http://www.softberry.com/) intended for annotation of plant genomes was used.

Softberry, Inc. is known as a leading developer of software tools for genomic research focused on computational methods of high throughput biomedical data analysis, including software to support next-generation sequencing technologies, transcriptome analysis (with RNASeq data), SNP detection and selection of disease-specific SNP subsets.

NSITE-PL service was used to search for promoter/functional motives (http://www.softberry.com/berry.phtml?topic=nsitep&group=programs&subgroup=promoter) (Shahmuradov and Solovyev, 2015; Solovyev et al., 2010)

Annotation of Sub-Cellular Localization

ProtComp v. 9.0 service was used to study the localization of the studied gene products throughout cell compartments.

RESULTS AND DISCUSSION

Despite both the fundamental knowledge gained from relevant studies concerning the wheat genome and the importance of the crop, a comprehensive genome-wide analysis of gene content was not conducted until recently (Dabab Nahas et al., 2019). This was because of the large size, repeat content, and polyploid complexity of the genome. However, assembly of the 17-Gb allohexaploid genome of *Triticum aestivum* faced major difficulties, because it is composed of three large, repetitive, and closely related genomes. In the current study, the Dreb genes of bread wheat were analyzed *in silico*. For this purpose, four genes encoding the Dreb genes were selected from the GenBank database, and information about them was obtained. The first information we analyzed was Access Nr. JQ004969.1, the beta isoform of the DREB AP2 binding factor, labeled "*Triticum aestivum* DREB AP2 binding factor beta isoform mRNA, complete cds". The length of this mRNT is 1286 bp. The protein-encoding area is shown to be smaller, in other words, it is located in the area between 64-264 nucleotides. The id of the protein corresponding to this gene: "AEX59145.1". This gene is translated to the amino acid sequence as follows:
"MTVDRKHAEAAAAAPFEIPALQPGRTCGA
EESTRSHVLVKPIGKSDLGDHVMGLIQSLKR
SGDGKK".

The second researched information is an AP2-containing protein-encoding gene labeled "*Triticum aestivum* AP2-containing protein (Dreb1) mRNA, complete cds" with Access Nr. AF303376.1. The length of this locus is 1292 bp. The region encoding the protein covers an area between 252-1088 nucleotides. The id of the protein corresponding to this gene in GenBank is: "AAL01124.1". The translation of the gene into an amino acid sequence was as follows: "METGGSKREGDCPGQERKKKVRRRSTGPD SVAETIKKWKEENQKLQQENGSRKAPAKGS KKGCMAGKGGPENSNCAYRGVRQRTWGK WVAEIREPNRGNRLWLGSFPTAVEAARAYD DAARAMYGAKARVNFSEQSPDANSGCTLA PPLPMSNGATAASHPSDGKDESESPPSLISNA PTAALHRSDAKDESESAGTVARKVKKEVSN DLRSTHEEHKTLEVSQPKGKALHKAANVSY DYFNVEEVLDMIIVELSADVKMEAHEEYQD GDDGFSLFSY".

The Dreb gene searched in NCBI under the name "*Triticum aestivum* Wdreb2 mRNA for EREBP / AP2 type transcription factor, complete cds" is information available under Accession Nr. AB193608.1. The length of the current locus was 1456 bp. The region encoding the protein covers the area between 123-1157 nucleotides. The amino acid sequence of the EREBP / AP2 type transcription factor protein corresponding to this gene is shown as follows:

```
CLUSTAL format alignment by MAFFT (v7.481)

AAL01124.1       METGGSKREGDC----------PGQERKKKVRRRSTGPDSVAETIKKWKEENQKLQQE--
ABA08424.1       METGGSKREGDC----------PGQERKKKVRRRSTGPDSVAETIKKWKEENQKLQQE--
BAD97369.1       MTVDRKHAEAAAAAPFEIPALQPG--RKKRPRRSRDGPNSVSETIRRWKEVNQQLEHDPQ
                 * .. .: *. .          **  ***: **   **:**:***::*** **:*:::

AAL01124.1       --NGSRKAPAKGSKKGCMAGKGGPENSNCAYRGVRQRTWGKWVAEIREPNRGNRLWLGSF
ABA8424.1        --NGSRKAPAKGSKKGCMAGKGGPENSNCAYRGVRQRTWGKWVAEIREPNRGNRLWLGSF
BD97369.1        GAKRARKPPAKGSKKGCMLGKGGPENTQCGFRGVRQRTWGKWVAEIREPNRVSRLWLGTF
                  : :**.********** *******::*.:******************** .*****:*

AAL01124.1       PTAVEAARAYDDAARAMYGAKARVNFSEQSPDANSGCTLAPPLPMSNGATAASHP-SDGK
BA0424.1         PTAVEAARAYDDAARAMYGAKARVNFSEQSPDANSGCTLAPPLPMSNGATAASHP-SDGK
BAD97369.1       PTAEDAARAYDEAARAMYGALARTNFPVHPAQAPAVAVPAAIEGVVRGASASCESTTTST
                 *** :******:******** **.**. :..:* : .. *.    : .**:*:... : ..
AAL01124.1       DESESPPSLISNAPTAALH-RSDAKDESESAGTVARKVKKEVSNDL---------RSTHE
AB08424.1        DESESPPSLISNAPTAALH-RSDAKDESESAGTVARKVKKEVSNDL---------RSTHE
AD97369.1        NHSDVASSLPRQAQAPEIYSQPDALESTES--VVLESVEHYSHQDTVPDAGSSISRSTSE
                 :.*: ..**  :* :. :: :.** :.:**  .* ..*::    :*         *** *

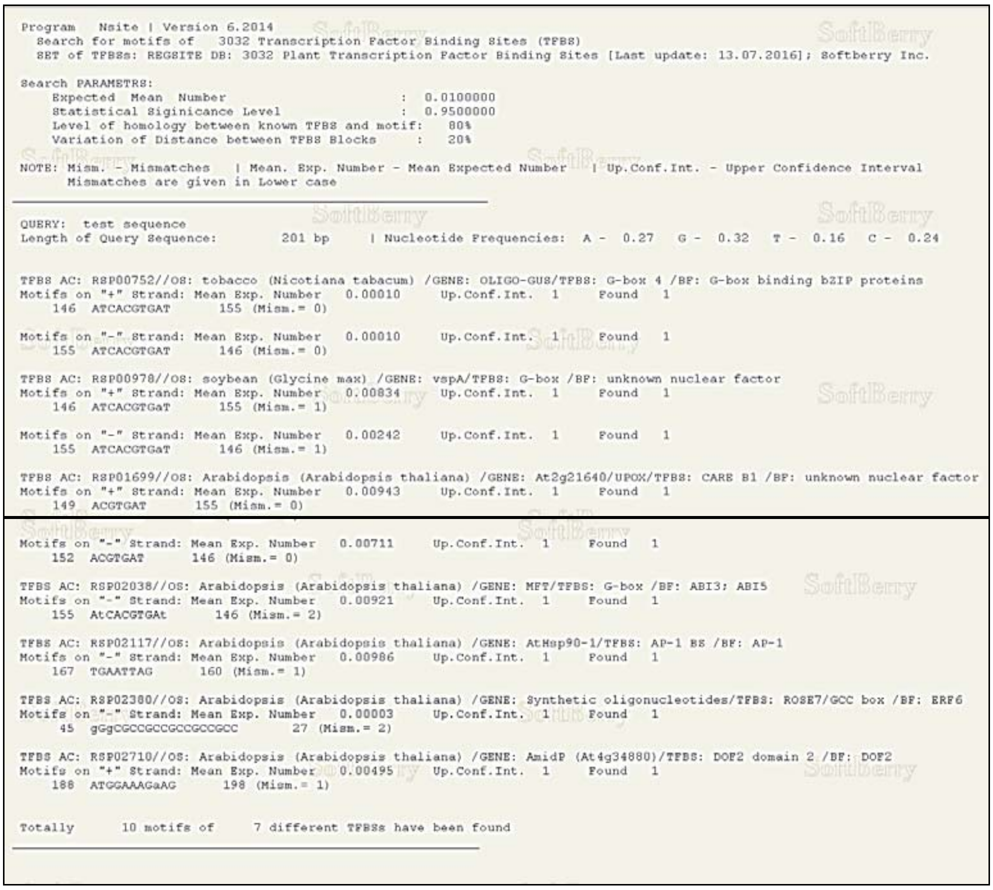
AAL01124.1       EH--KTLE--VSQPKGKALHKAANVSYDYFNVEEVL-----DMIIVEL----------SA
AA08424.1        EH--KTLE--VSQPKGKALHKAANVSYDYFNVEEVL-----DMIIVEL----------SA
BAD97369.1       EDVFEPLEPISSLPDGEA---------DGFDIEELLRLMEADPIEVELVTGGSWNGGANT
                *.  :.**   * *.*:*         * *::**:*     * * ***          .:

AAL01124.1       DVKMEAHEE-YQDGDDGFSL-------FSY---------------------------
ABA08424.1       DVKMEAHEE-YQDGDDGFSL-------FSY---------------------------
BAD97369.1       GVEMGQQEPLYLDGLDQGMLEGMLQSDYPYPMWISEDRAMHNSAFHDAEMSEFFEGL
                 .*:*  :*  * ** *   *       :.*
```

Fig. 1. CLUSTAL format alignment by MAFFT. Specific nuclear localization signal (NLS) area for the Dreb genes is highlighted in green, the amino acid substitutions in this area are highlighted in red. The valine 14 and glutamic acid 19 amino acids of the AP2 domain are yellow. The specific sequence for α-helix is pink, and the amino acid changes observed in this region are highlighted in blue.

"MTVDRKHAEAAAAAPFEIPALQPGRKKRP RRSRDGPNSVSETIRRWKEVNQQLEHDPQG AKRARKPPAKGSKKGCMLGKGGPENTQCG FRGVRQRTWGKWVAIREPNRVSRLWLGTFP TAEDAARAYDEAARAMYGALARTNFPVHP AQAPAVAVPAAIEGVVRGASASCESTTTSTN HSDVASSLPRQAQAPEIYSQPDALESTESVV LESVEHYSHQDTVPDAGSSISRSTSEEDVFEP LEPISSLPDGEADGFDIEELLRLMEADPIEVE LVTGGSWNGGANTGVEMGQQEPLYLDGLD QGMLEGMLQSDYPYPMWISEDRAMHNSAF HDAEMSEFFEGL". The protein corresponding to this gene is placed in GenBank with the id "BAD97369.1".

Finally, the last gene we studied is information placed under the name "*Triticum aestivum* genome D dehydration-responsive element-binding protein (Dreb1) gene, complete cds" with Accession Nr. DQ195068.1. This locus is longer and amounts to 1748 bp. The region encoding the protein consists of areas between nucleotides 20-69 and 771-1557, and covers two exons. The id of this protein, which combines the element responsible for dehydration, is "ABA08424.1" in GenBank. The translation of a nucleotide sequence into an amino acid sequence was as follows: "METGGSKREGDCPGQERKKKVRRRSTGPD SVAETIKKWKEENQKLQQENGSRKAPAKGS KKGCMAGKKGGPENSNCAYRGVRQRTWGK WVAEIREPNRGNRLWLGSFPTAVEAARAYD DAARAMYGAKARVNFSEQSPDANSGCTLA PPLPMSNGATAASHPSDGKDESESPPSLISNA PTAALHRSDAKDESESAGTVARKVKKEVSN DLRSTHEEHKTLEVSQPKGKALHKAANVSY DYFNVEEVLDMIIVELSADVKMEAHEEYQD GDDGFSLF"



```
Program   Nsite | Version 6.2014
  Search for motifs of   3032 Transcription Factor Binding Sites (TFBS)
  SET of TFBSs: REGSITE DB: 3032 Plant Transcription Factor Binding Sites [Last update: 13.07.2016]; Softberry Inc.

Search PARAMETRS:
    Expected  Mean  Number                       :  0.0100000
    Statistical Siginicance Level                :  0.9500000
    Level of homology between known TFBS and motif:   80%
    Variation of Distance between TFBS Blocks     :   20%

NOTE: Mism. - Mismatches   | Mean. Exp. Number - Mean Expected Number   | Up.Conf.Int. - Upper Confidence Interval
      Mismatches are given in Lower case

QUERY:  test sequence
Length of Query Sequence:        201 bp    | Nucleotide Frequencies:  A -  0.27   G -  0.32   T -  0.16   C -  0.24

TFBS AC: RSP00752//OS: tobacco (Nicotiana tabacum) /GENE: OLIGO-GUS/TFBS: G-box 4 /BF: G-box binding bZIP proteins
Motifs on "+" Strand: Mean Exp. Number   0.00010     Up.Conf.Int.  1     Found   1
    146  ATCACGTGAT      155 (Mism.= 0)

Motifs on "-" Strand: Mean Exp. Number   0.00010     Up.Conf.Int.  1     Found   1
    155  ATCACGTGAT      146 (Mism.= 0)

TFBS AC: RSP00978//OS: soybean (Glycine max) /GENE: vspA/TFBS: G-box /BF: unknown nuclear factor
Motifs on "+" Strand: Mean Exp. Number   0.00834     Up.Conf.Int.  1     Found   1
    146  ATCACGTGaT      155 (Mism.= 1)

Motifs on "-" Strand: Mean Exp. Number   0.00242     Up.Conf.Int.  1     Found   1
    155  ATCACGTGaT      146 (Mism.= 1)

TFBS AC: RSP01699//OS: Arabidopsis (Arabidopsis thaliana) /GENE: At2g21640/UPOX/TFBS: CARE B1 /BF: unknown nuclear factor
Motifs on "+" Strand: Mean Exp. Number   0.00943     Up.Conf.Int.  1     Found   1
    149  ACGTGAT      155 (Mism.= 0)

Motifs on "-" Strand: Mean Exp. Number   0.00711     Up.Conf.Int.  1     Found   1
    152  ACGTGAT      146 (Mism.= 0)

TFBS AC: RSP02038//OS: Arabidopsis (Arabidopsis thaliana) /GENE: MFT/TFBS: G-box /BF: ABI3; ABI5
Motifs on "-" Strand: Mean Exp. Number   0.00921     Up.Conf.Int.  1     Found   1
    155  AtCACGTGAt      146 (Mism.= 2)

TFBS AC: RSP02117//OS: Arabidopsis (Arabidopsis thaliana) /GENE: AtHsp90-1/TFBS: AP-1 BS /BF: AP-1
Motifs on "-" Strand: Mean Exp. Number   0.00986     Up.Conf.Int.  1     Found   1
    167  TGAATTAG      160 (Mism.= 1)

TFBS AC: RSP02380//OS: Arabidopsis (Arabidopsis thaliana) /GENE: Synthetic oligonucleotides/TFBS: ROSE7/GCC box /BF: ERF6
Motifs on "-" Strand: Mean Exp. Number   0.00003     Up.Conf.Int.  1     Found   1
     45  gGgCGCCGCCGCCGCC      27 (Mism.= 2)

TFBS AC: RSP02710//OS: Arabidopsis (Arabidopsis thaliana) /GENE: AmidP (At4g34880)/TFBS: DOF2 domain 2 /BF: DOF2
Motifs on "+" Strand: Mean Exp. Number   0.00495     Up.Conf.Int.  1     Found   1
    188  ATGGAAAGaAG      198 (Mism.= 1)

Totally     10 motifs of     7 different TFBSs have been found
```

Fig. 2. Annotation of functional motifs for *Triticum aestivum* DREB AP2 binding factor gen (Accession Nr. JQ004969.1) using NSITE-PL (http://www.softberry.com).

```
5724 multiple located sequences are accepted
ProtComp Version 9.0. Identifying sub-cellular location (Plant)
Seq name:  test sequence, Length=278
Significant similarity in Location DB - Nuclear
Database sequence: AC=Q0JQF7 Location:Nuclear  DE  Dehydration-responsive element-
Score=98, Sequence length=275, Alignment length=270
Predicted by Neural Nets - Extracellular (Secreted) with score    1.0
Integral Prediction of protein location: Nuclear with score    8.9
Location weights:     LocDB / PotLocDB / Neural Nets / Pentamers / Integral
 Nuclear               10.0 /      3.0 /        0.00 /      0.06 /     8.93
 Plasma membrane        0.0 /      0.0 /        0.97 /      0.03 /     0.71
 Extracellular          0.0 /      0.0 /        0.97 /      1.02 /     0.00
 Cytoplasmic            0.0 /      0.0 /        0.00 /      2.14 /     0.00
 Mitochondrial          0.0 /      0.0 /        0.00 /      1.43 /     0.00
 Endoplasm. retic.      0.0 /      0.0 /        0.00 /      0.09 /     0.00
 Peroxisomal            0.0 /      0.0 /        0.97 /      0.00 /     0.10
 Golgi                  0.0 /      0.0 /        0.10 /      0.31 /     0.00
 Chloroplast            0.0 /      0.0 /        0.00 /      0.21 /     0.01
 Vacuolar               0.0 /      0.0 /        0.00 /      0.00 /     0.26
```

Fig. 3. Annotation of sub-cellular localization for *Triticum aestivum* genome D dehydration-responsive element-binding protein (protein_id “BAD97369.1”) using ProtComp 9.0 (http://www.softberry.com).

CLUSTAL format alignment by MAFFT (v7.481) revealed significantly conservative areas in the amino acid sequences of these genes. The beta isoform (protein_id: "AEX59145.1") of the DREB AP2 binding factor, with Accession Nr. JQ004969.1 in GenBank has been excluded due to inconsistencies in this alignment. Homology is more pronounced between the Dreb1 gene with Accession Nr. AF303376.1 (protein_id. "AAL01124.1") and Wdreb2 genes with Accession Nr. AB193608.1 in GenBank. In the amino acid sequences of both proteins, a specific nuclear localization signal (NLS) for the Dreb genes is observed in the peptide sequence (RKKKVR) (highlighted in green in Figure 1). In the amino acid sequence encoded by the dehydration-responsive element-binding protein (Dreb1) gene (Accession Nr. DQ195068.1) of the *Triticum aestivum* genome D, in the 4th position of the signal peptide sequence (RKKRPR), arginine is located instead of lysine and in the 5th position, proline amino acid is located instead of valine (substituted amino acids in Figure 1 are highlighted red).

Besides, alignment of amino acid sequences corresponding to the Dreb gene by MAFFT revealed the AP2 domain with the two conserved functional amino acids (valine (V) and glutamic acid (E)) at the 14th and 19th residues which play crucial roles in recognition of the DNA - binding sequence (Fig. 1). However, according to the results of further research, E19 might not be as necessary as V14 for this case (Sakuma et al. 2002). Nevertheless, both of these amino acids are found in the three proteins we studied (highlighted yellow in the figure).

The characteristic sequence for the α-helix (VEAARAYDDAARAMYG) was also identified in our analysis *in silico*. Interestingly, BAD97369.1 protein contains amino acid substitutions also in this area. Valine, the primary amino acid involved in the formation of α-helix, was replaced by glutamine in this protein, and the second glutamine was replaced by the amino acid aspartate. In the ninth position of this domain, on the contrary, aspartate was replaced by glutamine. It should be noted that the gene of this protein, which changes both in the NLS sequence and in the α-helix region, is a Dreb gene specific to the D genome. For the first time, Pandey et al. (2014) built the tertiary structure of DREB2 protein from wheat by homology modeling based on the crystal structure of GCC-box binding domain of Arabidopsis thaliana, which contributed to understanding the structure-function relationships. Protein docking with the DNA containing GCC-box revealed more similarities between AP2/EREBP protein of A. thaliana and T. aestivum. A protein was found to interact through their β-sheet, with the major DNA groove by hydrogen and hydrophobic bond, which provides structural stability to the molecule. This model comprises a three-stranded antiparallel β-

heet followed by α-helix and relatively unstructured C'-terminal.

To search for functional motives in the selected DREB genes, NSITE-PL service for annotation of plant genomes of the Softberry, Inc. (http://www.softberry.com/) software was used (Shahmuradov and Solovyev, 2015). Ten motifs for seven different transcription factor-binding sites (TFBS) were found in *Triticum aestivum* DREB AP2 binding factor beta isoform with Accession Nr. JQ004969.1 (Figure 2). Seventeen motifs for 15 different TFBS were found in *Triticum aestivum* AP2-containing protein (Dreb1) with Accession Nr. AF303376.1. Nine motifs for eight different TFBS in *Triticum aestivum* Wdreb2 mRNA for EREBP/AP2 type transcription factor (Accession Nr. AB193608.1) and seventeen motifs for fifteen different TFBS in *Triticum aestivum* genome D dehydration-responsive element-binding protein (Dreb1) gene (Accession Nr. DQ195068.1) were found.

Annotation of sub-cellular localization was performed by the ProtComp v. 9.0 service of Softberry for the studied products. Nuclear location was determined for three of the transcription factors studied (Figure 3), and extracellular localization was found only for the DREB AP2 binding factor beta isoform. This result requires more in-depth research.

In silico identification and characterization of the genes in various organisms under different conditions got importance due to growing data in the data bases (Dabab Nahas et al., 2019). Our analyses could be a scientific base to understand Dreb genes and proteins to further wet lab studies in wheat plants.

ACKNOWLEDGMENT

This work was supported by the Science Development Foundation under the President of the Republic of Azerbaijan – Grant № EIF-ETL-2020-2(36)-16/15/3-M-15).

REFERENCES

Alotaibi F., Alharbi S., Alotaibi M., Al Mosallam M., Motawei M., Alrajhi A. (2021) Wheat omics: Classical breeding to new breeding technologies. *Saudi journal of biological sciences*, **28(2):** 1433.

Dabab Nahas L., Al-Husein N., Lababidi G., Hamwieh A. (2019) In-silico prediction of novel genes responsive to drought and salinity stress tolerance in bread wheat (*Triticum aestivum*). *Plos one*, **14(10):** e0223962.

Hassan S., Berk K., Aronsson H. (2021) Evolution and identification of DREB transcription factors in the wheat genome: modeling, docking and simulation of DREB proteins associated with salt stress. *Journal of Biomolecular Structure and Dynamics*, p. 1-14. doi: 10.1080/07391102.2021.1894980

Niu X., Luo T., Zhao H., Su Y., Ji W., Li H. (2020) Identification of wheat DREB genes and functional characterization of TaDREB3 in response to abiotic stresses. *Gene*, **740:** 144514.

Pandey B., Sharma P., Saini M., Pandey D. M., Sharma I. (2014) Isolation and characterization of dehydration-responsive element-binding factor 2 (DREB2) from Indian wheat (*Triticum aestivum* L.) cultivars. *Australian Journal of Crop Science*, **8(1):** 44.

Rathan N. D., Sehgal D., Thiyagarajan K., Singh R. P., Singh A. M., Velu G. (2021) Identification of genetic loci and candidate genes related to grain zinc and iron concentration using a zinc-enriched wheat 'Zinc-Shakti'. *Frontiers in Genetics*, **12:** Article ID 652653. doi: 10.3389/fgene.2021.652653

Sakuma Y., Liu Q., Dubouzet J. G., Abe H., Shinozaki K., Yamaguchi-Shinozaki K. (2002) DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration-and cold-inducible gene expression. *Biochemical and biophysical research communications*, **290(3):** 998-1009.

Shahmuradov I. A., Solovyev V. V. (2015) Nsite, NsiteH and NsiteM computer tools for studying transcription regulatory elements. *Bioinformatics*, **31(21):** 3544-3545.

Shahzad R., Shakra Jamil S. A., Nisar A., Amina Z., Saleem S., Iqbal, M. Z., ... Wang X. (2021) Harnessing the potential of plant transcription factors in developing climate resilient crops to improve global food security: Current and future perspectives. *Saudi Journal of Biological Sciences*, **28(4):** 2323.

Solovyev V.V., Shahmuradov I.A., Salamov A.A. (2010) Identification of promoter regions and regulatory sites. In: *Computational biology*

of transcription factor binding. Humana Press, Totowa, NJ. pp. 57-83.

Walkowiak S., Gao L., Monat C., Haberer G., Kassa M.T., Brinton J., ... Pozniak C.J. (2020) Multiple wheat genomes reveal global variation in modern breeding. *Nature*, **588:** 277–283.

Zhu T., Wang L., Rimbert H., Rodriguez J.C., Deal K.R., De Oliveira R., ... Luo M.C. (2021) Optical maps refine the bread wheat *Triticum aestivum* cv. Chinese Spring genome assembly. *The Plant Journal*, doi: 10.1111/tpj.15289. Online ahead of print.

Yumşaq buğdada Dreb transkripsiya faktoru genlərinin *in silico* analizi

S. M. Rüstəmova

AMEA-nın Molekulyar Biologiya və Biotexnologiyalar İnstitutu, Bakı, Azərbaycan

Buğdada quraqlıqla induksiya olunan transkriptomda əsas yer tutan AP2/ERF transkripsiya faktorları ailəsinin üzvlərindən dehidratasiyaya cavabdeh element birləşdirən (Dreb) genlər *in silico* analiz edilmişdir. Bu məqsədlə NCBI məlumat bazasından genlər (qeydiyyat nömrələri: AF303376.1, AB193608.1, KM520370.1, DQ195068.1) haqqında məlumatlar əldə edilmişdir. Hər bir genə uyğun proteinlərin FASTA məlumatları MAFFT CLUSTAL format düzlənmə proqramı ilə müqayisəli analiz edilərək əsas konservativ sahələr müəyyən edilmişdir. Tədqiq olunan genlərin hamısında AP2 domen üçün spesifik iki konservativ funksional amin turşusu - valin və qlutamin 14-cü və 19-cu vəziyyətlərdə müəyyən edilmişdir. D genomu üçün spesifik olan dehidratasiyaya cavabdeh elementi birləşdirən proteində (DQ195068.1) nüvədə lokalizasiya siqnalının (NLS) və fəza quruluşunda α-spiral quruluşun yaranmasında iştirak edən sahələrdə spesifik amin turşu əvəzlənmələri müəyyən edilmişdir. Əldə olunan nəticələr buğda bitkisində Dreb genlərin gələcək laboratoriya tədqiqatları üçün elmi əsas ola bilər.

Açar sözlər: *Dreb, AP2 domen, nüvədə lokalizasiya siqnalı (NLS), α-spiral quruluş, in silico analiz*

In silico анализ генов факторов транскрипции Dreb в мягкой пшенице

С. М. Рустамова

Институт молекулярной биологии и биотехнологий НАН Азербайджана, Баку, Азербайджан

С использованием методов *in silico* были изучены относящиеся к представителям семейства факторов транскрипции AP2/ERF и играющие ключевую роль в индуцированной засухой транскриптоме пшеницы гены, связывающие элемент, ответственный за дегидратацию (Dreb). С этой целью информация о соответствующих генах (номер доступа AF303376.1, AB193608.1, KM520370.1, DQ195068.1) была получена из NCBI. Данные FASTA белков, соответствующих каждому гену, были сравнительно проанализированы с помощью программного обеспечения для выравнивания MAFFT CLUSTAL, идентифицированы основные консервативные области чтения. Две консервативные функциональные аминокислоты, специфичные для домена AP2 - валин и глутамин - идентифицированы в положениях 14 и 19 во всех изученных генах. Определенные аминокислотные замены были идентифицированы в белке (DQ195068.1), который связывается с элементом дегидратации, специфичным для генома D, в областях, участвующих в формировании сигнала ядерной локализации (NLS) и структуры α-спирали. Полученные результаты могут стать научной основой для будущих лабораторных исследований генов Dreb у пшеницы.

Ключевые слова: *Dreb, домен AP2, сигнал ядерной локализации (NLS), α-спираль, in silico анализ*